

EXHIBIT 2





Nucleotide

PubMed	Nucleotide	Protein	Genome	Structure	PopSet	Taxonomy	OMIM	
Search	Nucleotide	for					Go	Clear
Limits		Index		History		Clipboard		
Display	Default View	as	HTML	Save	Add to Clipboard			

☐ 1: X02750 **Human liver** PubMed, Protein, Related Sequences, Taxonomy, OMIM, LinkOut
mRNA for
protein C

```

LOCUS       HSPROTC                1843 bp      mRNA                      PRI          05-APR-1995
DEFINITION  Human liver mRNA for protein C.
ACCESSION   X02750
VERSION     X02750.1   GI:35689
KEYWORDS    protein C; signal peptide.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1843)
AUTHORS     Beckmann,R.J., Schmidt,R.J., Santerre,R.F., Plutzky,J.,
            Crabtree,G.R. and Long,G.L.
TITLE       The structure and evolution of a 461 amino acid human protein C
            precursor and its messenger RNA, based upon the DNA sequence of
            cloned human liver cDNAs
JOURNAL     Nucleic Acids Res. 13 (14), 5233-5247 (1985)
MEDLINE     85269639
COMMENT     Data kindly reviewed (27-MAR-1986) by G. Long.
FEATURES             Location/Qualifiers
     source          1..1843
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS           98..1483
                     /note="protein C precursor"
                     /codon_start=1
                     /protein_id="CAA26528.1"
                     /db_xref="GI:763120"
                     /db_xref="SWISS-PROT:P04070"
                     /translation="MWQLTSLLLFVATWGISGTPAPLDSVFSSSERAHQVLRIRKRAN
SFLEELRHSSLERECIEEICDFEEAKEIFQNVDDTLAFWSKHVDGDQCLVLPLEHPCA
SLCCGHTCIDGIGSFSCDCRSGWEGRFCQREVSVFLNCSLDNGGCTHYCLEEVGWRRR
SCAPGYKLDDLLQCHPAVKFPCGRPWKRMEKKRSHLKRDTEDQEDQVDPRLIDGKMT
RRGDSWPQVVLVLDSSKKKLACGAVLIHPSWVLTAAHCMDESKLLVRLGEYDLRRWEKW
ELDLDIKEVFVHPNYSKSTTDNDIALHLHAQPATLSQTIVPICLPDSGLAERELNQAG
QETLVGTGWGYHSSREKEAKRNRTFVLNFIKIPVVPHNECSEVMSNMVSENMLCAGILG
DRQDACEGDSGGPMVASFHGTWFLVGLVSWGEGCGLLHNYGVYTKVSRYLWDWIHGHIR
DKEAPQKSWAP"
     sig_peptide    98..196
     misc feature    197..223
                     /note="propeptide (aa -9 to -1)"
     misc feature    224..358
                     /note="gamma carboxylation domain (aa 1-45)"
     misc feature    224..688
                     /note="light chain (aa 1-155)"
     misc feature    359..496
                     /note="EGF-domain I (aa 46-91)"
     misc feature    497..634
                     /note="EGF-domain II (aa 92-137)"
     misc feature    653..730

```

misc feature /note="activation peptide region (aa 144-169)"
695..1531
misc feature /note="heavy chain (aa 57-419)"
731..1531
misc feature /note="serine protease region (aa 170-419)"
1759..1764
misc feature /note="polyA signal"

BASE COUNT 417 a 530 c 564 g 332 t

ORIGIN

```
1 ctgcagggggg gggggggggg gggggctgtc atggcggcag gacggcgaac ttgcagtatc
61 tccacgaccc gccctacag gtgccagtgc ctccagaatg tggcagctca caagcctcct
121 gctgttcgtg gccacctggg gaatttcggg cacaccagct cctcttgact cagtgttctc
181 cagcagcgag cgtgcccacc aggtgctgcg gatccgcaaa cgtgccaaact ccttcctgga
241 ggagctccgt cacagcagcc tggagcggga gtgcatagag gagatctgtg acttcgagga
301 ggccaaggaa attttccaaa atgtggatga cacactggcc ttctggtcca agcacgtcga
361 cggtgaccag tgcttgggtc tgcccttggg gcacccgtgc gccagcctgt gctgcgggca
421 cggcacgtgc atcgacggca tcggcagctt cagctgcgac tgccgcagcg gctgggaggg
481 ccgcttctgc cagcgcgagg tgagcttctt caattgctcg ctggacaacg gcggctgcac
541 gcattactgc ctagaggagg tgggctggcg gcgctgtagc tgtgcgcctg gctacaagct
601 gggggacgac ctctgcagt gtcaccccg cagtgaagttc ccttgtggga ggccctggaa
661 gcggatggag aagaagcgca gtcacctgaa acgagacaca gaagaccaag aagaccaagt
721 agatccgcgg ctcatgtatg ggaagatgac caggcgggga gacagccctt ggcaggtggt
781 cctgctggac tcaaagaaga agctggcctg cggggcagtg ctcatccacc cctcctgggt
841 gctgacagcg gccactgca tggatgagtc caagaagctc cttgtcaggc ttggagagta
901 tgacctgcgg cgctgggaga agtgggagct ggacctggac atcaaggagg tcttcgtcca
961 cccaactac agcaagagca ccaccgacaa tgacatcgca ctgctgcacc tggcccagcc
1021 cgccaccctc tcgcagacca tagtgcccat ctgcctcccg gacagcggcc ttgcagagcg
1081 cgagctcaat caggccggcc aggagaccct cgtgacgggc tggggctacc acagcagccg
1141 agagaaggag gccaagagaa accgcacctt cgtcctcaac ttcataaaga tccccgtggt
1201 cccgcacaat gagtgcagcg aggtcatgag caacatggtg tctgagaaca tgctgtgtgc
1261 gggcatcctc ggggaccggc aggatgcctg cgagggcgac agtggggggc ccatggtcgc
1321 ctcttccac ggcacctggt tcttggtggg cctggtgagc tggggtgagg gctgtgggct
1381 ccttcacaac tacggcgttt acaccaaagt cagccgctac ctcgactgga tccatgggca
1441 catcagagac aaggaagccc ccagaagag ctgggcacct tagcgacctt ccctgcaggg
1501 ctgggctttt gcatggcaat ggatgggaca ttaaagggac atgtaacaag cacaccggcc
1561 tgctgttctg tccttccatc cctcttttgg gctcttctgg agggaaagta catttactga
1621 gcacctgttg tatgtcacat gccttatgaa tagaatctta actcctagag caactctgtg
1681 ggggtggggag gagcagatcc aagttttgcg ggggtctaaag ctgtgtgtgt tgagggggat
1741 actctgttta tgaaaaagaa taaaaaacac aaccacgaaa aaaaaaaaaa aaaaaaaaaa
1801 aaaaaaaaaa aaaaaaaccc ccccccgccc cccccctg cag
```

//

Restrictions on Use | Write to the HelpDesk
NCBI | NLM | NIH